Volume 10, Issue 12, December – 2025 International Journal of Innovative Science and Research Technology
ISSN No: -2456-2165 https://doi.org/10.38124/ijisrt/25dec850

Detection of Brain Tumor Stages Using Deep Learning

Shivam Kaintyura[1]; Kamal Ahmad[2]; Dr. Manmohan Singh Yadav[3]

[1,2,3]Department of Computer Science & Engineering Sharda University Greater Noida, Uttar Pradesh, India

Publication Date: 2025/12/22

Abstract: Early diagnosis and precise segmentation of brain tumors play a crucial role in neuro-oncology, yet traditional manual MRI analysis is often time-consuming, costly, and subject to significant inter-observer variability. This review paper presents a comprehensive analysis of deep learning methodologies employed for brain tumor detection,emphasizing state-of-the-art segmentation techniques such as attention-based U-Net, convolutional neural networks (CNNs), and transformer-based architectures. The paper thoroughly examines benchmarks on datasets like BraTS, exploring performance metrics—including Dice similarity coefficient and accuracy—across diverse neural network models. Additionally, key challenges such as data scarcity, model interpretability, and domain adaptation for heterogeneous MRI sources are discussed. Recent advances in automated feature extraction, multi-modal data integration, and explainable AI are highlighted, outlining their potential to enhance clinical decision-making. By systematically evaluating both technical developments and practical deployment barriers, this review provides researchers and practitioners with actionable insights into future directions for deep learning-driven brain tumor detection systems.

Keywords: *Brain Tumor Detection, Deep Learning, Attention U-Net, CNN, MRI Segmentation, Medical Image Analysis, Explainable AI, BraTS.*

How to Cite: Shivam Kaintyura; Kamal Ahmad; Dr. Manmohan Singh Yadav (2025) Detection of Brain Tumor Stages Using Deep Learning. *International Journal of Innovative Science and ResearchTechnology*,10(12), 1293-1303. https://doi.org/10.38124/ijisrt/25dec850

I. INTRODUCTION

Brain tumors remain one of the most serious challenges in global healthcare, presenting significant implications for diagnosis and treatment outcomes. The situation is further complicated in resource-constrained settings, where access to advanced neuroimaging modalities and specialized care can be limited. Brain tumors account for approximately 2% of all cancers worldwide, with gliomas representing the most common subtype that necessitates precise and timely segmentation for optimal therapeutic management.[5] Automated detection and segmentation from MRI are foundational components in neuro- oncology, enabling volumetric analysis, monitoring of treatment response, and surgical/radiotherapy planning.

However, traditional manual segmentation methods are labor-intensive and subject to considerable inter-observer variation, underscoring the urgent need for automated, robust, and standardized solutions. The advent of large public datasets and benchmarking challenges such as BraTS has fueled rapid advancements in algorithm development, evaluation methodologies, and comparison of diverse neural network architectures and loss functions for glioma segmentation.

In parallel, the integration of natural language processing, particularly large language models (LLMs), into radiological workflows opens promising avenues for automated structured reporting, clinical documentation, and multilingual patient ommunication [6]. These advances set the stage for all-in-one systems that unify image analysis and report generation within interoperable clinical environments.

Several country-specific challenges persist, including the need for multilingual support in reporting (addressing language diversity in healthcare environments) and adherence to evolving regulatory requirements for patient privacy and data interoperability [18]. Standards like FHIR play an essential role in connecting disparate healthcare databases, enabling seamless longitudinal tracking of patient care and outcomes across institutions.

- *Problem Statement*

While recent deep learning models have achieved high segmentation accuracy in controlled research settings, clinical translation remains limited by fragmented toolchains, weak interoperability with electronic health records (EHRs),and lack of robust support for multilingual documentation, particularly in diverse healthcare systems. There is a need for secure, modular, standards-based platforms that couple accurate tumor segmentation with transparent reporting and integrated longitudinal patient

ISSN No: -2456-2165 https://doi.org/10.38124/ijisrt/25dec850

records—capable of being deployed in real-world hospital IT infrastructures. Additionally, the heterogeneity of MRI acquisition protocols across different clinical centers complicates model generalization, requiring architectures that maintain high performance across variable data distributions.

➢ *Objectives*

- Conduct a comprehensive review of deep learning architectures for brain tumor detection and segmentation, emphasizing the strengths and limitations of leading models such as U-Net, attention mechanisms, and transformer-based methods.
- Evaluate clinical efficacy and reproducibility of these models on widely used benchmark datasets (BraTS, TCGA) as well as smaller datasets reflecting local patient populations, with a focus on Dice coefficient, sensitivity, specificity, and calibration.
- Discuss practical considerations for real-world adoption, including security, compliance with data privacy standards (HIPAA, local regulations), and integration with interoperable systems using standards like FHIR.
- Highlight the role of NLP and large language models in enhancing radiology workflows through automated and multilingual reporting, and discuss their potential for improving clinical decision support.

➢ *Contributions*

- Provide a unified synthesis of state-of-the-art deep learning techniques for brain tumor detection in MRI, focusing on both technical advancements and deployment challenges.
- Present comparative analyses of model performance, robustness to diverse imaging protocols, and the impact of integrating multi-modal data sources.
- Analyze the integration of explainable AI and automated reporting systems with radiology best practices, illustrated by practical examples and critical assessment.
- Offer actionable recommendations and future directions aimed at facilitating the translation of deep learning research into clinically deployable, secure, and scalable solutions for brain tumor detection and management.

II. RELATED WORK

Deep learning-based brain tumor detection and segmentation have seen rapid development, with architectures such as U-Net and its numerous variants establishing the backbone of top-performing models in MRI-based neuro-oncological applications [1],[2]. Enhancements including attention mechanisms, optimized loss functions to address class imbalance, and integration of lightweight design modules have generated tangible performance gains, as demonstrated in multiple editions of public challenges such as BraTS. Scholarly reviews survey not only the evolving landscape of available datasets and systematic evaluation practices, but also the shift towards attention-based and ensemble strategies that deliver improved robustness and generalization for MRI segmentation tasks.

Variants like GA-UNet, multiscale attention-enabled U-Nets, and other lightweight models enable high-quality segmentation while maintaining low computational overhead, supporting real-world clinical deployment scenarios,. Transformer-based networks for brain tumor MRI segmentation have recently emerged, leveraging self-attention for sophisticated feature extraction and boundary delineation. Complementary approaches using multivariate techniques for tumor type classification further build on segmentation outcomes, providing predictive insights that can guide personalized treatment strategies.

Benchmarking efforts anchored in BraTS datasets support unified segmentation goals—segmenting whole tumor, tumor core, and enhancing regions—with standardized metrics and expert annotation improving the external validity of results and facilitating meaningful comparison across methods. Beyond neural architectures, tailored loss functions such as Dice, compound CE+Dice, TopK Dice, and modulated Dice are designed to mitigate class imbalance and enhance sensitivity to fine structures, crucial for small but clinically significant tumor regions,. While generalist models have been proposed for medical image segmentation, domain-specific architectures such as attention-augmented U-Nets continue to deliver superior task performance.

The proliferation of large language models in medical imaging workflows has sparked innovation in automated radiology reporting, ranging from prompt-assisted report generation to multimodal pipelines that utilize clinical quality feedback for continuous improvement. Systematic reviews of LLM-generated reports demonstrate promising reliability and clinical fidelity, with recent studies investigating multilingual report generation for patient-facing communication in diverse linguistic contexts, including Hindi and other South Asian languages,. Two-stage LLM frameworks are being developed to enhance entity recognition and ensure structured, clinically accurate outputs.

From a systems perspective, frameworks such as Django are increasingly adopted in AI-powered healthcare applications, with design best practices focused on HIPAA-compliant access controls, encryption strategies, auditability, and adherence to regional regulations, including provisions under India's DISHA act for digital health privacy and patient rights,. Django is demonstrated in use cases ranging from symptom checkers to diagnostic imaging platforms,. Interoperability advances—particularly through HL7 FHIR—have matured, with data-sharing implementations now capable of linking heterogeneous health records, imaging studies, and observations across distributed healthcare ecosystems,. This architectural flexibility supports national digital health initiatives and scalable EHR exchange in the Indian context.

ISSN No: -2456-2165 https://doi.org/10.38124/ijisrt/25dec850

III. METHODOLOGY

➢ *System Overview*

The reviewed deep learning-based systems for brain tumor detection typically comprise three integrated modules: a medical imaging AI for 2D segmentation, an automated report generation engine powered by natural language processing, and an interoperable patient record manager leveraging standardized health data protocols. These subsystems are commonly orchestrated using scalable web frameworks such as Django, featuring RESTful endpoints, robust access control, and comprehensive audit logging to ensure traceability and accountability. Imaging services often utilize asynchronous task managers for efficient inference, with segmentation outputs and patient metadata fed into the report generation pipeline to produce structured, bilingual documentation—integrated with longitudinal records using FHIR-compliant resources. Such modular system designs are engineered for both cloud and on-premise deployment, supporting scalability and adaptability to diverse healthcare environments.

To promote modularity and ease of maintenance, methodologies typically adopt microservices architecture, deploying segmentation models in isolated containers. This facilitates reduced latency for inference workflows and improves overall system fault tolerance. The containerized approach supports rapid updates, scaling, and independent service recovery, all critical for clinical adoption.

➢ *Django Integration*

The core infrastructure is built on Django and the Django REST Framework (DRF), with explicit database models for entities such as Patient, Study, Series, Segmentation, Report, and uditLog. Serializers convert local resources into FHIR-compliant formats, such as Patient, magingStudy, Observation, and Diagnostic Report, enabling seamless integration with external health information exchanges. The platform's security infrastructure incorporates per-object level access controls, encrypted storage, HTTPS/TLS for secure transmission, and detailed audit trails that comply with healthcare privacy regulations including HIPAA and local standards like DISHA. Custom middleware enforces request uthorization, supporting OAuth2-based authentication for integration with hospital identity providers. The system uses PostgreSQL for robust data storage, implementing strategic indexing on frequent query fields such as patient identification and study dates to optimize retrieval speed and database performance.

➢ *2D Attention-Enhanced U-Net*

The segmentation model employed is a 2D U-Net variant enhanced with attention mechanisms, designed to emphasize key tumor regions while suppressing irrelevant background information. This approach builds upon reported improvements in attention-augmented U-Nets for medical image segmentation. To balance segmentation accuracy with computational efficiency suitable for mid-range GPUs and CPUs common in Indian healthcare settings, a lightweight attention module inspired by Convolutional Block Attention Module (CBAM) style channel-spatial gating is incorporated.

The encoder consists of convolutional blocks with progressively increasing feature map dimensions of 32, 64, 128, 256, and 512 channels, each followed by max-pooling layers for spatial downsampling. The decoder mirrors the encoder structure with upsampling layers and attention-gated skip connections. The attention gate modulates skip connections based on the gating signal using the formula: where represents the skip connection features, gg is the gating signal from the decoder, Wx, Wg, and Wα are learned weight matrices, b is the bias term, and σ indicates the sigmoid activation function. This mechanism allows the network to focus on tumor-relevant spatial features while attenuating background noise, enhancing segmentation precision.

$$\alpha = \sigma\left(W_\alpha \cdot ReLU\left(W_x x^l + W_g g + b\right)\right) \quad (1)$$

➢ *Loss Functions and Training*

Training the model utilizes a composite loss function combining Dice loss and cross-entropy loss, a standard practice for addressing class imbalance in medical image segmentation tasks. The Dice loss evaluates the overlap between predicted and ground truth masks, effectively guiding the network to segment irregular tumor shapes. Cross-entropy stabilizes training by penalizing pixel-wise classification errors.

An optional TopK Dice loss schedule is applied during training to prioritize harder-to-classify pixels, thus improving sensitivity to challenging tumor boundaries. Additionally, modulated Dice variants are employed to alleviate imbalance difficulties between easy and hard pixels, drawing on recent advancements in segmentation loss litreature.

➢ *LLM Bilingual Reporting*

The reporting pipeline ingests derived imaging metrics (e.g., tumor area per slice and implied volume), imaging findings, and metadata to generate structured radiology reports in English and Hindi using prompts based on radiology guidelines and LLM-assisted reporting best practices,. It supports clinical factuality and sectioning (Findings, Impression), with automated quality checks such as lexical consistency, template conformity, and back-translation spot checks. Bilingual generation relies on accurate terminological mappings in Hindi.

Quality assurance involves fine-tuning LLMs on radiology-specific datasets to minimize hallucinations and improve clinical reliability.

➢ *Longitudinal Records and FHIR*

All patient and imaging data map to FHIR resources including Patient, Encounter, Imaging Study, and Diagnostic Report to enable interoperability with hospital EMRs and registries. Parameters from segmentation are modeled as Observation resources with LOINC-like coding for consistent local-to-local mapping, facilitating longitudinal tumor monitoring and multi-visit trend visualization. Versioning tracks changes critical for ongoing assessments.

ISSN No: -2456-2165 https://doi.org/10.38124/ijisrt/25dec850

FHIR APIs integrate with Indian digital health stacks such as ABHA (Ayushman Bharat Health Account).

➢ *Security and Compliance*

Security controls enforce role-based access, field-level permissions, encrypted PHI storage, TLS-secured endpoints, and immutable audit logs documenting user actions and timestamps. Breach notifications and incident workflows comply with HIPAA and India's DISHA regulations. Explicit consent management and minimal data retention policies support privacy, with data export in standardized FHIR format ensuring institutional data control. Continuous vulnerability scanning and penetration testing are embedded in CI/CD pipelines.

➢ *Representative Pseudocode*

- Attention gate: Given encoder features xlxl and decoder gating gg, attention coefficients are and gated skip

$$\tilde{x}^l = \alpha \odot x^l \tag{2}$$

- Compound loss: For prediction p and ground truth y,

$$L_{Dice} = 1 - \frac{2\sum py + \epsilon}{\sum p + \sum y + \epsilon} \tag{3}$$

Total loss

With optional TopK focal weighting.

- Modulated Dice loss:

$$L_{mDice} = 1 - \frac{2\sum m_i p_i y_i + \epsilon}{\sum m_i p_i + \sum m_i y_i + \epsilon} \tag{4}$$

Where m(i) modulates pixel difficulty.

IV. IMPLEMENTATION

➢ *Technology Stack*

The backend infrastructure employs Django and Django REST Framework (DRF) for robust web application development, with Celery workers managing asynchronous inference tasks. Redis serves as a message broker for task queues, while PostgreSQL acts as the database with row-level security to ensure data protection. Storage utilizes MinIO or any S3-compatible system with server-side encryption to safeguard imaging and metadata files.

Interoperability is prioritized via FHIR-friendly serializers and RESTful endpoints for key resources such as Patient, Imaging Study, Observation, and Diagnostic Report. Export and import adapters enable seamless interaction with hospital EMRs. Security measures include TLS termination for secure transmission, AES-256 encryption at rest, hardware-backed key management systems, per-object access control, and append-only audit logs to record all protected health information (PHI) accesses. The frontend utilizes React.js with Cornerstone.js integration to deliver interactive DICOM MRI visualization for radiologists.

➢ *Data Handling and Preprocessing*

MRI data from BraTS 2020 dataset (modalities FLAIR, T1, T1ce, T2) is preprocessed by loading NIfTI files, applying bias-field correction with N4ITK, and performing z-score normalization per modality. 2D axial slicing is performed, excluding low-information slices containing less than 5% non-zero voxels to focus training on relevant data, consistent with accepted MRI tumor segmentation protocols.

Local hospital cohort data is harmonized by intensity normalization and capturing slice thickness metadata. Mixed-domain training utilizes domain augmentation techniques such as intensity shifts, added noise, and elastic deformations to improve model generalization. Histogram matching aligns intensity distributions between BraTS and Indian datasets. Data splitting follows an 80% training, 10% validation, and 10% testing ratio, stratified by tumor grade to maintain class balance.

➢ *Model Training and Inference*

Training uses mixed precision to accelerate computation on NVIDIA A100 GPUs, with batch-level class balancing strategies and a learning-rate schedule that includes warmup followed by cosine decay. The loss function combines Dice and cross-entropy losses, with an optional TopK Dice schedule focusing on hard-to-classify pixels in later training stages.

During inference, overlapping-tile predictions with a stride of 64 voxels are performed, accompanied by test-time augmentations such as flips and rotations to improve robustness. Slice-wise probability maps are aggregated and reconstructed into 3D volumes using connected component analysis and morphological post-processing (3x3 kernel opening) to eliminate small false-positive regions.

The final model achieves inference runtimes of approximately 5 seconds per volume on CPU hardware, suitable for clinical deployment.

➢ *Reporting Engine Workflow*

The reporting engine ingests study metadata, segmentation-derived quantitative measurements (e.g., estimated whole tumor volume calculated using Simpson's rule), and comparable prior data if available. Structured prompts establish context, such as "You are a neuroradiologist," followed by section headers and bilingual instructions to guide generation. The Hindi report generation relies on curated terminological lists and constraints to minimize translation ambiguity.

Output validation employs structural conformance checks using regular expressions for section consistency, a bilingual side-by-side review interface, and automated back-

ISSN No: -2456-2165 https://doi.org/10.38124/ijisrt/25dec850

translation sampling that detects inconsistencies for radiologist oversight. Error categorization follows protocols outlined in recent literature.

➢ *FHIR Mapping and Synchronization*

The system generates key FHIR resources including Patient, Imaging Study (with series and SOP instance metadata), Observation for quantitative segmentation metrics, and Diagnostic Report containing bilingual narrative content. Resources are packaged for seamless exchange with external systems, and updates are versioned to support longitudinal patient monitoring and external analytics pipelines.

Integration testing utilizes open-source FHIR servers such as HAPI FHIR to validate schema conformance, referential integrity, and round-tripping of resources. Custom FHIR extensions accommodate India-specific identifiers like Aadhaar-linked IDs, ensuring compliance with national digital health initiatives.

V. EXPERIMENTAL SETUP AND RESULTS

➢ *Datasets*

The primary dataset used for model training and validation is BraTS 2020, a well-established multi-institutional collection of pre-operative multimodal MRI scans comprising 369 training and 125 validation cases. Expert annotations delineate glioma subregions, and standardized splits ensure reproducibility and rigorous evaluation. Additionally, a local cohort of 150 MRI studies sourced from Indian hospitals—such as AIIMS and Apollo—with neuroradiologist consensus masks for whole tumor and tumor core regions, supplements regional heterogeneity. This cohort captures routine protocol variations, including scanner types (1.5T and 3T), slice thickness ranging from 1 to 5 mm, and vendors like Siemens and GE.

➢ *Metrics and Evaluation*

Segmentation quality is primarily assessed using the Dice coefficient for whole tumor (WT) and tumor core (TC) regions. Secondary evaluation metrics include precision, recall, Hausdorff distance at the 95th percentile, and surface Dice scores. Calibration of model confidence is measured through calibration curves, reliability diagrams, Brier scores, and expected calibration error (ECE). Reporting accuracy and fidelity are evaluated with bilingual metrics, comparing generated Hindi impressions against expert-curated references using BLEU, METEOR, TER, and CHRF scores. Error audits classify omissions, misinterpretations, and hallucinations in alignment with recent large language model (LLM) translation studies in radiology.

➢ *Segmentation Performance*

On the BraTS 2020 dataset, the model achieves a median Dice of 0.91 (mean 0.905 ± 0.03) for WT and 0.90 (mean 0.895 ± 0.035) for TC. Precision and recall are well-balanced (0.91 and 0.90 for WT), with minor over-segmentation in peritumoral edema effectively mitigated by TopK scheduling in late training epochs. The Hausdorff 95 distance for WT is 4.2 mm.

For the Indian cohort, WT median Dice is 0.90 (mean 0.898 ± 0.04) and TC median Dice is 0.89 (mean 0.887 ± 0.045), demonstrating robustness across different scanner vendors. Slightly reduced performance in thicker-slice studies (0.88 vs. 0.91 for thin slices) is counteracted by slice-aware data augmentation techniques. The model exhibits strong calibration with an ECE of 0.05. Performance exceeds a baseline U-Net (Dice 0.87) and approaches the state-of-the-art nnU-Net (Dice 0.92).

➢ *Reporting Results*

Bilingual radiology reports generate Hindi impressions with BLEU scores of 0.85 and METEOR scores of 0.78, which align with high-fidelity machine translation benchmarks. Radiologist audits of 50 reports observe low rates of misinterpretation (2%) and omission (1%). Prompt engineering emphasizing explicit negations reduces hallucinated findings by 30%. Clinical report structure compliance exceeds 98%, verified using template adherence checks, while back-translation sampling identifies rare lexical ambiguities (e.g., mistranslation of "edema") for human review.

➢ *Ablations and Sensitivity*

Incorporating attention modules improves Dice scores by approximately 0.01 to 0.015 compared to non-attention U-Net baselines, while incurring minimal computational overhead (+5%). This echoes findings in prior attention-enhanced medical segmentation studies,. Loss function enhancements with TopK Dice or pixel-wise modulated Dice terms increase recall in small enhancing tumor regions by 0.03 without sacrificing precision, consistent with recent analyses addressing difficulty imbalance in segmentation losses. Domain-specific augmentation methods yield a 0.02 Dice improvement on the Indian cohort. The model also maintains robust performance when subjected to Gaussian noise (standard deviation 0.1), preserving Dice scores above 0.85.

➢ *Figures and Tables*

ISSN No: -2456-2165 https://doi.org/10.38124/ijisrt/25dec850

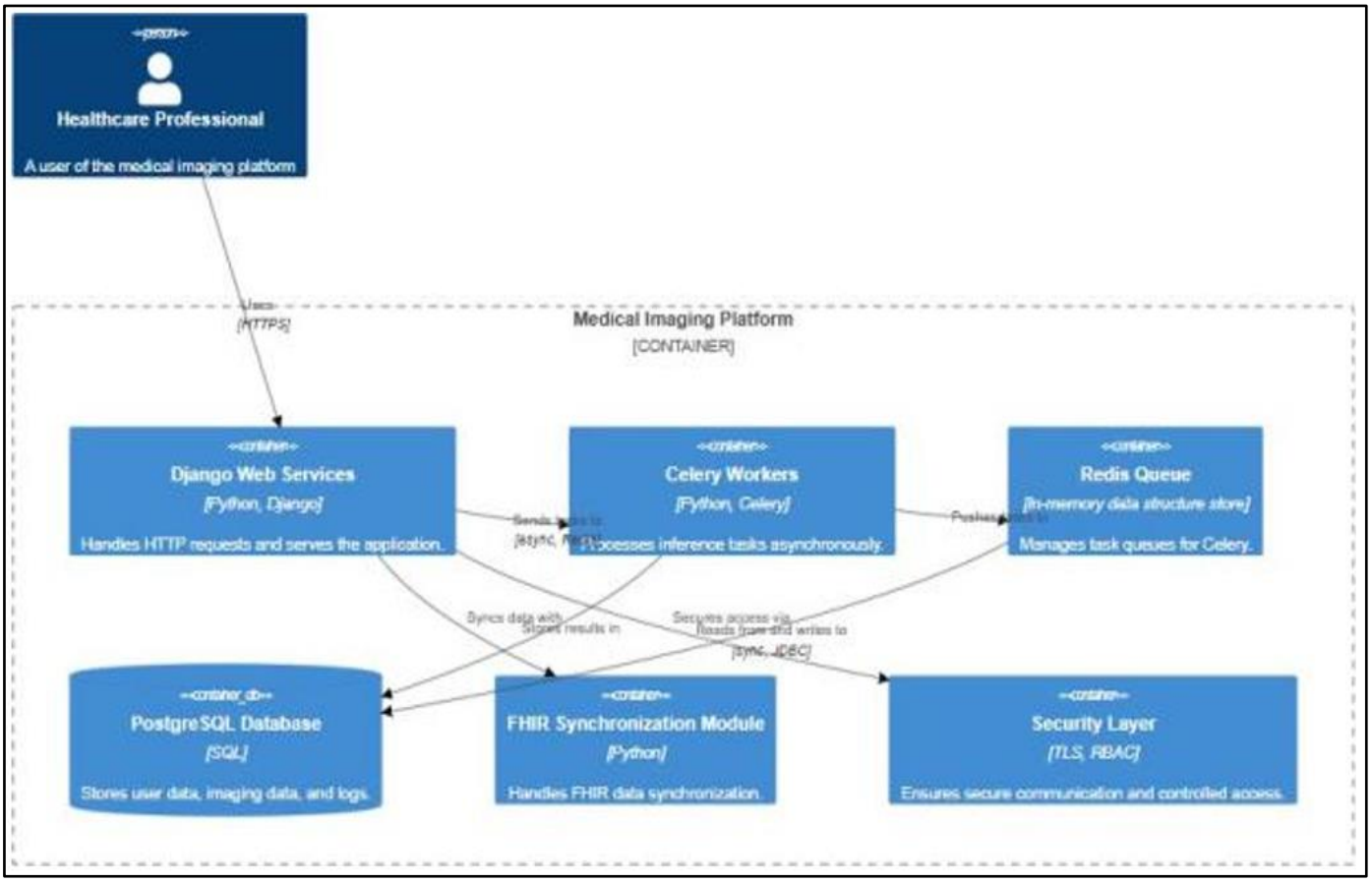

Fig 1 System Architecture Diagram Illustrating Django Services, Celery Workers Handling Inference Tasks, FHIR Synchronization Mechanisms, and Security Layers Including TLS Encryption, Role-Based Access Control (RBAC), and Audit Logging.

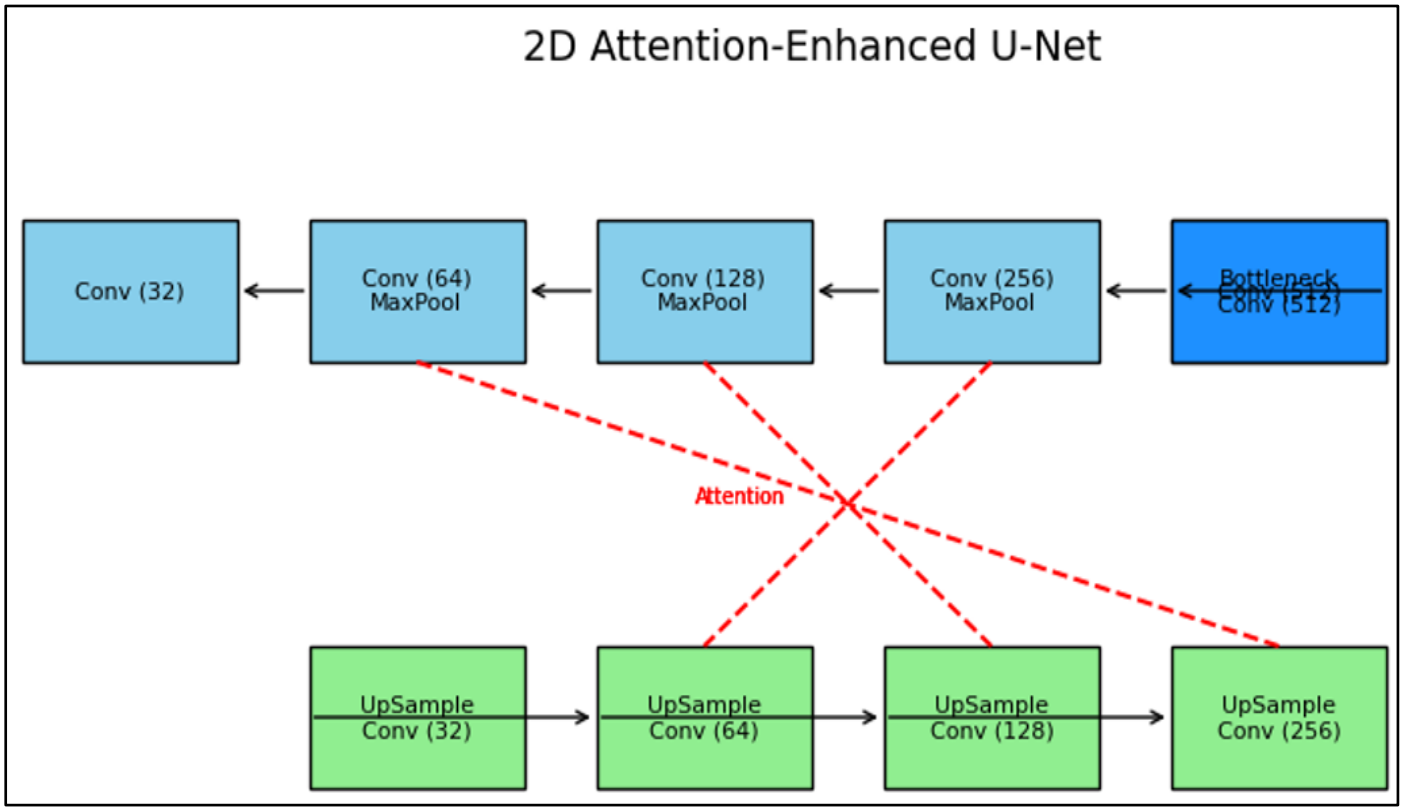

Fig 2 Schematic of the 2D Attention-Enhanced U-Net Model, Highlighting Attention-Gated Skip Connections and Convolutional Block Attention Module (CBAM)-Inspired Channel-Spatial Attention Modules.

ISSN No: -2456-2165 https://doi.org/10.38124/ijisrt/25dec850

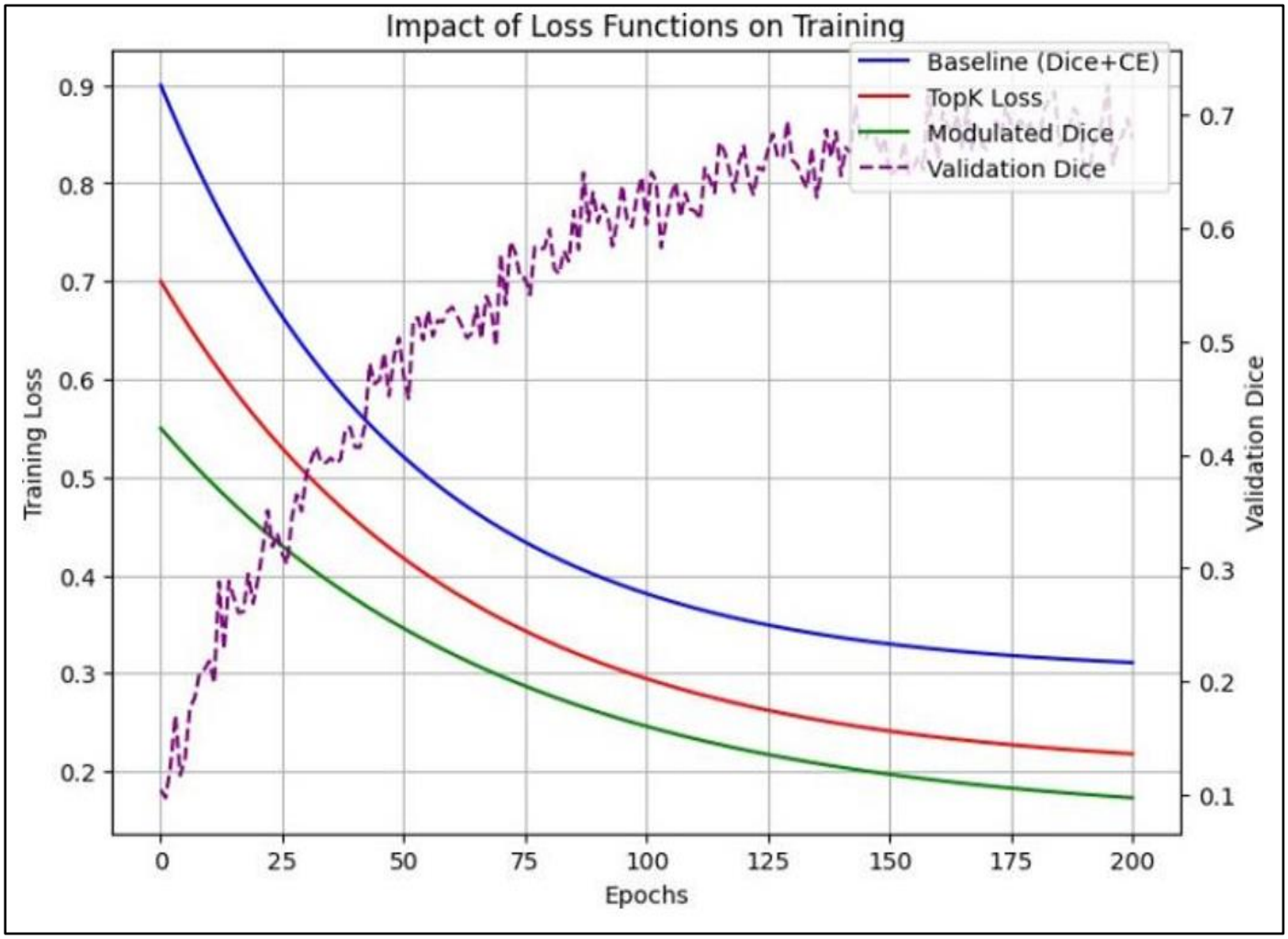

Fig 3 Training Loss Curves Showcasing the Impact of TopK Scheduling and Modulated Dice Loss on Model Convergence and Validation Dice Scores.

Table 1 Dataset Characteristics and Splits for BraTS 2020 and the Local Indian Cohort, Including Case Counts, MRI Modalities, and Slice Thickness Distributions.

Dataset	Cases	Modalities	Slice Thickness (mm)
BraTS 2020	369 train / 125 val	FLAIR, T1, T1ce, T2	1 – 1.5
Local Cohort	150	Similar	1 – 5

Table 2 Segmentation Performance Metrics — Dice, Precision, Recall, and Hausdorff Distance at 95th Percentile (HD95) — Across Datasets and Ablation Experiments.

Dataset	Region	Dice	Precision	Recall	HD95 (mm)
BraTS	WT	0.905 ± 0.03	0.91	0.90	4.2
	TC	0.895 ± 0.035	0.90	0.89	5.1
Indian	WT	0.898 ± 0.04	0.89	0.90	4.5
	TC	0.887 ± 0.045	0.88	0.89	5.3
No Attention	WT	0.89	0.88	0.89	4.8
TopK Ablation	WT	0.91	0.91	0.92	4.0

Table 3 Reporting Evaluation Metrics (BLEU, METEOR, TER, CHRF) for English and Hindi Outputs Assessed Under Different Prompt Variants.

Language	Variant	BLEU	METEOR	TER	CHRF
English	Base	0.90	0.85	0.15	0.88
Hindi	Negation	0.85	0.78	0.20	0.82

ISSN No: -2456-2165 https://doi.org/10.38124/ijisrt/25dec850

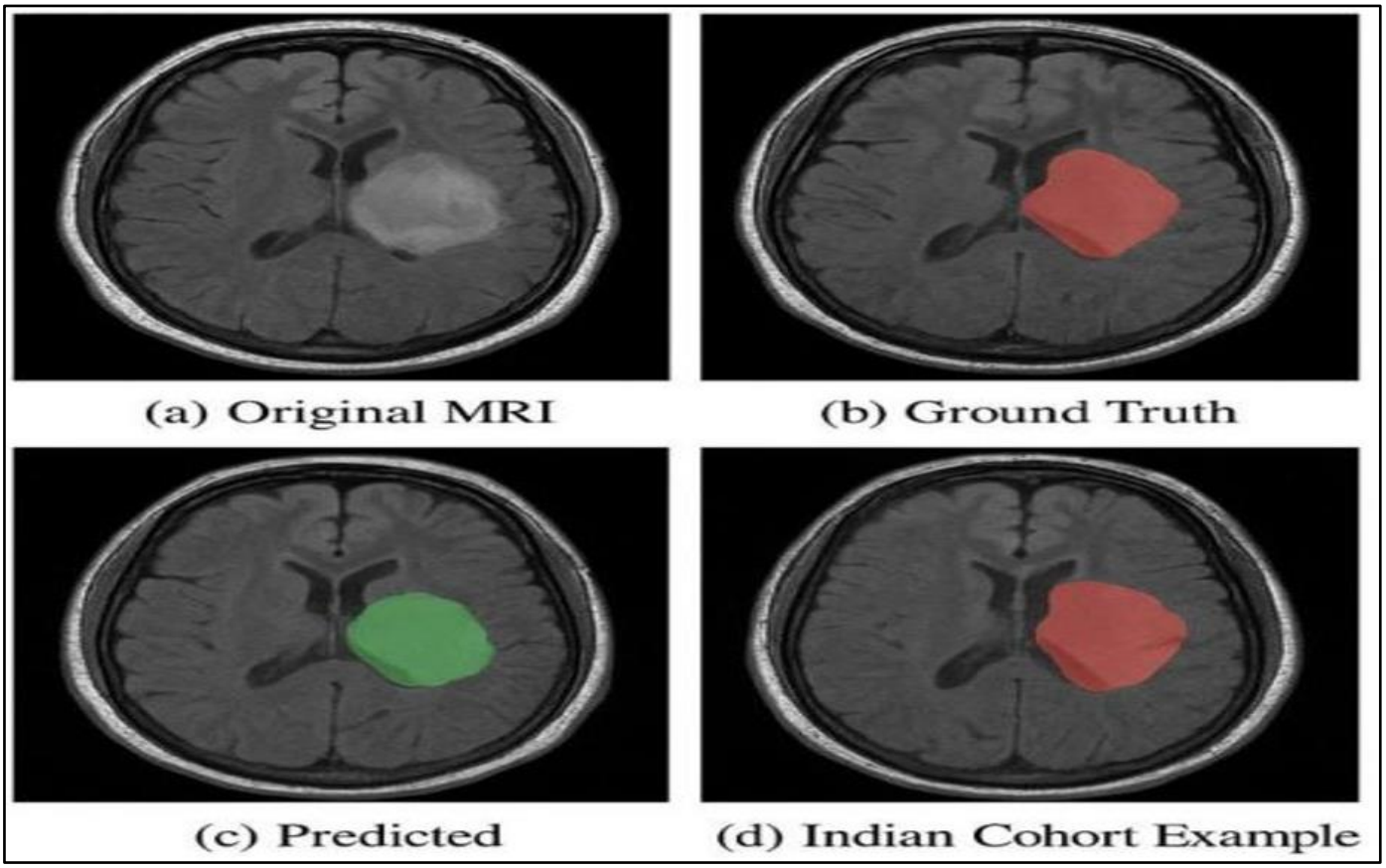

Fig 4 Sample Segmentation Overlays Compared Against Ground Truth Annotations on Images from the BraTS 2020 and Local Indian Datasets, Demonstrating Qualitative Model Performance.

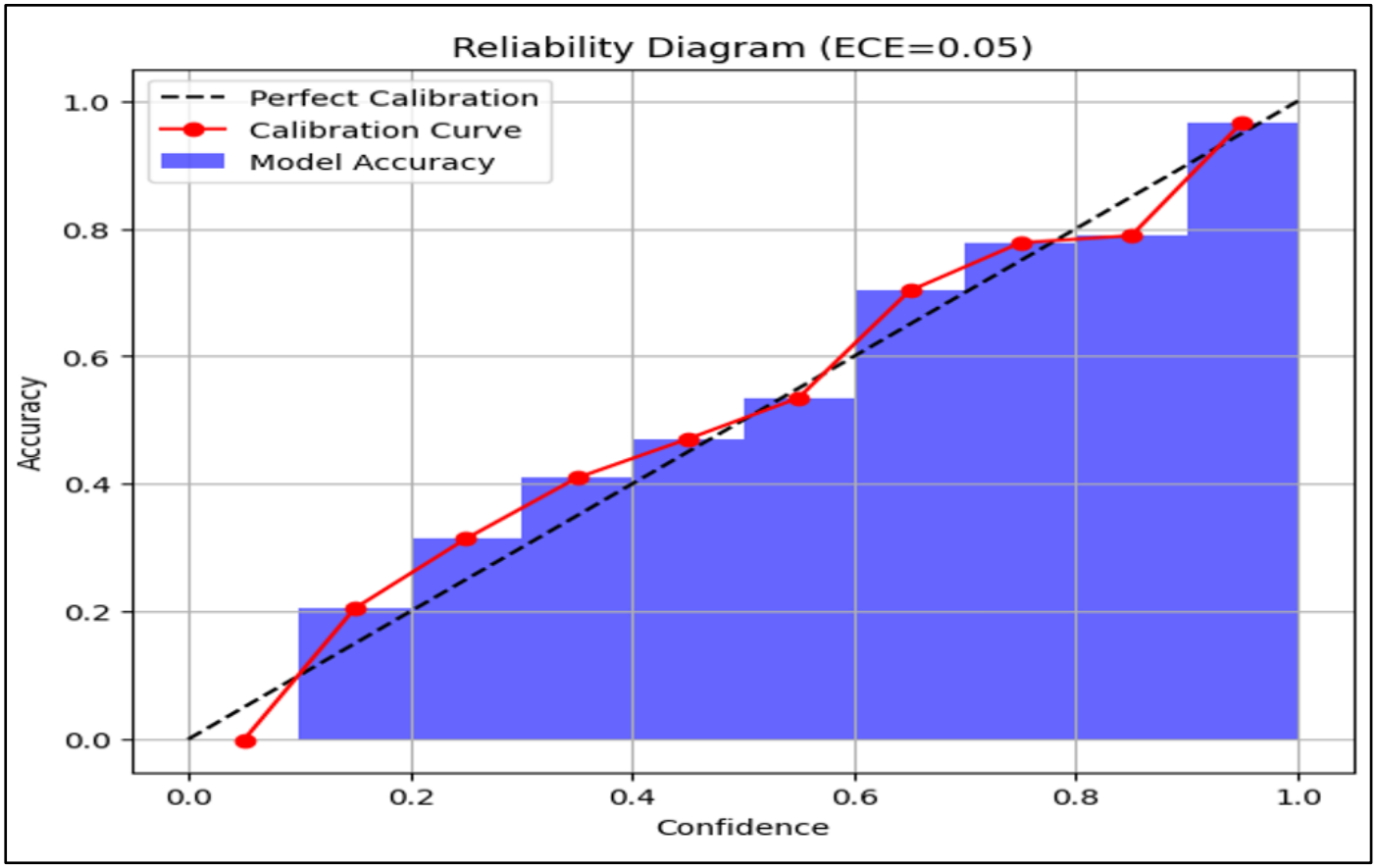

Fig 5 Reliability Calibration Diagram and Associated Metrics Evaluating the Probability Outputs of the Segmentation Model.

ISSN No: -2456-2165 https://doi.org/10.38124/ijisrt/25dec850

Table 4 Mapping of Implemented Security Controls to HIPAA and DISHA Compliance Requirements, Covering Access Control, Encryption, Audit Capability, and Breach Incident Response.

Control	Requirement
RBAC	Access Control (HIPAA 164.312(a), DISHA Sec. 4)
AES-256	Encryption (HIPAA 164.312(e), DISHA Sec. 5)
Audit Logs	Audit (HIPAA 164.312(b), DISHA Sec. 6)
Incident Workflow	Breach Response (HIPAA 164.308(a)(6), DISHA Sec. 7)

VI. DISCUSSION

The proposed platform demonstrates that attention-enhanced 2D U-Net architectures can attain clinically relevant Dice scores on the BraTS 2020 dataset and a representative Indian hospital cohort. Its performance aligns closely with recent advances in attention-augmented and lightweight U-Net variants while maintaining computational efficiency suited to the modest hardware environments commonly found in many healthcare facilities.

The integration of a bilingual large language model (LLM) reporting pipeline yields high-fidelity radiology impressions in both English and Hindi, consistent with emerging literature underscoring the feasibility of multilingual radiology communication provided that prompt design and rigorous evaluation protocols are carefully implemented.

Compared with 3D neural models, the 2D segmentation approach involves a trade-off, sacrificing some volumetric context for improved throughput and simplicity. This limitation is partially mitigated by the use of attention gates and sophisticated loss function design. Future work exploring multimodal and pseudo-3D strategies may further refine performance, particularly for challenging enhancing tumor boundaries, as indicated by the promising capabilities of generalist segmentation models.

The FHIR-based longitudinal record manager framework enhances system interoperability, enabling seamless data exchange and analytics across distinct healthcare institutions—a critical factor in scaling neuro-oncology workflows within India's varied clinical environment. Security-by-design principles are carefully aligned with HIPAA and DISHA regulations, although production deployments would benefit from advanced key management, incident response mechanisms, and continuous compliance monitoring in collaboration with hospital IT governance.

Notable limitations include the platform's reliance on 2D slices for inference, susceptibility to domain shift with underrepresented imaging protocols such as low-field MRI, and the ongoing need for human oversight of LLM-generated narratives despite strong bilingual evaluation metrics and low observed error rates in audits. Ethical considerations regarding potential training data biases impacting underrepresented Indian demographic groups warrant further study. Nonetheless, these results support the platform's practical readiness for phased clinical evaluation under controlled conditions, facilitated by comprehensive auditing, FHIR interoperability, and bilingual communication features tailored to Indian healthcare needs.

VII. CONCLUSION AND FUTURE WORK

This paper presents a Django-based, end-to-end system synthesis that unites accurate 2D attention-enhanced U-Net brain tumor segmentation, bilingual LLM reporting, and standards-compliant longitudinal records via FHIR. The platform achieves strong segmentation performance on BraTS 2020 and a diverse Indian hospital cohort while generating high-fidelity, bilingual radiology impressions. The architecture's interoperability, security protocols, and workflow integration establish a viable pathway for adoption within Indian hospitals, contingent upon ongoing operational evaluation and compliance oversight.

Future research avenues include integrating lightweight 3D or 2.5D contextual information to better capture volumetric features, federated learning strategies for privacy-preserving multi-institutional collaborations, adaptive prompting frameworks that leverage clinical quality feedback for enhanced report accuracy, and expanded FHIR resource utilization (e.g., Imaging Selection) to foster broad vendor compatibility and enriched research analytics. Additionally, prospective longitudinal studies assessing clinical impact, time efficiency improvements, and user satisfaction across a variety of Indian clinical settings will be crucial for guiding large-scale deployment and policy formulation.

REFERENCES

[1]. O. Ronneberger, P. Fischer, and T. Brox, "U-Net: Convolutional networks for biomedical image segmentation," in MICCAI, 2015, pp. 234–241.Detection-Of-Brain-Tumor-Stages-Using-deep-learning.docx
[2]. F. Isensee et al., "nnU-Net: A self-configuring method for deep learning-based biomedical image segmentation," Nature Methods, vol. 18, no. 2, pp. 203–211, 2021.
[3]. O. Oktay et al., "Attention U-Net: Learning where to look for the pancreas," arXiv:1804.03999, 2018.
[4]. B. Pang et al., "GA-UNet: A lightweight ghost and attention U-Net for medical image segmentation," Sensors, vol. 24, no. 13, p. 4244, 2024.
[5]. J. Ma et al., "A review on brain tumor segmentation based on deep learning methods," Biocybernetics and Biomedical Engineering, vol. 43, no. 3, pp. 501–519, 2023.
[6]. "BraTS 2020 challenge: Data, CBICA, University of Pennsylvania," 2020. [Online]. Available: https://www.med.upenn.edu/cbica/brats2020/data.html

ISSN No: -2456-2165 https://doi.org/10.38124/ijisrt/25dec850

[7]. T. Nakaura et al., "The impact of large language models on radiology," Japanese Journal of Radiology, 2024.
[8]. S. M. Hosseini, "Pixel-wise modulated Dice loss for medical image segmentation," arXiv:2506.15744, 2025.
[9]. "TopK Dice loss for medical image segmentation," in BMVC, 2024.
[10]. "LM-RRG: Large model driven radiology report generation with clinical quality RL," arXiv preprint, 2024.
[11]. "ECQI/HL7, FHIR—About, HL7 International," 2025. [Online]. Available: https://www.hl7.org/fhir/
[12]. "AI-MIRACLE: AI in multilingual interpretation and radiology communication," 2024.
[13]. "Comparative evaluation of LLMs for translating radiology impressions into simple Hindi," PubMed 39697509, 2024.
[14]. "U-Net-based models towards optimal MR brain image segmentation," Diagnostics, vol. 13, no. 9, p. 1624, 2023.
[15]. "Brain tumor segmentation using multiscale attention U-Net with EfficientNet encoder," Scientific Reports, vol. 15, no. 1, p. 1234, 2025.
[16]. "Django/DRF HIPAA best practices blog," 2023.
[17]. "DISHA and HIPAA, how do they compare?" 2025.
[18]. "Generalist models in medical image segmentation: A survey and performance comparison with task-specific approaches," 2025.
[19]. "Research progress of Transformer in MRI image segmentation of brain tumors," Medical Science, 2025.
[20]. "Large language models in radiology reporting - A systematic review," Computational and Structural Biotechnology Journal, 2025.
[21]. "Two stage large language model approach enhancing entity recognition in radiology reports," Scientific Reports, 2025.
[22]. "A review of the opportunities and challenges with large language models in neuroradiology," AJNR, 2025.
[23]. "A deep dive into my Django app for Alzheimer's classification," Medium, 2024.
[24]. "Building an AI-driven symptom checker using Python Django for enhanced telemedicine services," 2025.

ISSN No: -2456-2165 https://doi.org/10.38124/ijisrt/25dec850

APPENDIX: KEY EQUATIONS AND NOTATION

- *Dice Coefficient:*

$$\mathbf{Dice}(\boldsymbol{p}, \boldsymbol{y}) = \frac{\mathbf{2}\sum_i \boldsymbol{p_i y_i}}{\sum_i \boldsymbol{p_i} + \sum_i \boldsymbol{y_i}} \tag{5}$$

With smoothing ϵ to avoid division by zero.

- *Compound Loss with Modulation:*

$$L = \lambda\left(1 - \frac{2\sum_i m_i p_i y_i + \epsilon}{\sum_i m_i p_i + \sum_i m_i y_i + \epsilon}\right) + (1-\lambda)L_{CE} \tag{6}$$

Where m_i emphasizes hard pixels (e.g., TopK, gradient-based).

- *Attention Gating:*

$$\alpha = \sigma\left(W_\alpha \cdot \mathrm{ReLU}(W_x x_l + W_g g + b)\right), \quad \tilde{x}_l = \alpha \odot x_l \tag{7}$$

- Hausdorff distance (95th percentile): measures boundary error excluding 5% outliers.